

Data Science Quiz

10 questions about our first lesson: what data science is, the types of tasks, data and testing models.

Name: _____ Date: _____ Score: _____ / 10

Tick **ONE** answer for each question. Take your time and read every choice. Good luck!

1

Data science mixes three things. Which ones?

- ☐ A Drawing, music and sport
- ☐ B Programming, maths & statistics, and knowing the topic well
- ☐ C Typing fast, gaming and reading
- ☐ D Only computers

2

What does machine learning mean?

- ☐ A A robot that tidies your room
- ☐ B A computer that learns the rules by itself from lots of examples
- ☐ C A computer that only follows rules written by a person
- ☐ D A machine that builds other machines

3

You want to predict tomorrow's temperature in Dubai (a number, like 38.5 °C). Which type of task is it?

- ☐ A Regression
- ☐ B Classification
- ☐ C Clustering
- ☐ D Anomaly detection

4

An app looks at a photo and answers "cat" or "dog". Which type of task is it?

- ☐ A Regression
- ☐ B Association rules
- ☐ C Classification
- ☐ D Clustering

5

What is true about supervised learning?

- ☐ A The computer has no data at all
- ☐ B A teacher watches the computer all day
- ☐ C The computer makes groups without any answers
- ☐ D We already know the right answers in our examples, so we can check the model

6

A shop sorts its customers into groups with similar shopping habits, without knowing the groups in advance. This is...

- ☐ A Clustering
- ☐ B Regression
- ☐ C Classification
- ☐ D Deployment

7

We use this table to predict if a student passes. Which column is the target?

Hours studied	Sleep (h)	Attendance	Passed?
5	8	95%	Yes
1	5	60%	No

- ☐ A Hours studied
- ☐ B Sleep (h)
- ☐ C Passed?
- ☐ D Attendance

8

In a survey, the same column contains "Yes", "Y" and "1", all meaning yes. What problem is this?

- ☐ A Missing data
- ☐ B Inconsistency
- ☐ C Duplicate
- ☐ D Outlier

9

Why do we keep a test set aside?

- ☐ A To make the dataset bigger
- ☐ B To delete bad data
- ☐ C To train the model faster
- ☐ D To check the model on new data it has never seen

10

A model gets every answer right on its training data, but makes lots of mistakes on new data. This is called...

- ☐ A Overfitting
- ☐ B Underfitting
- ☐ C Clustering
- ☐ D A good fit

Finished? Check your answers once more, then send your quiz back to your teacher. Great job!